# Matrix Element Method for H->$\tau\tau$

Thomas Strebler

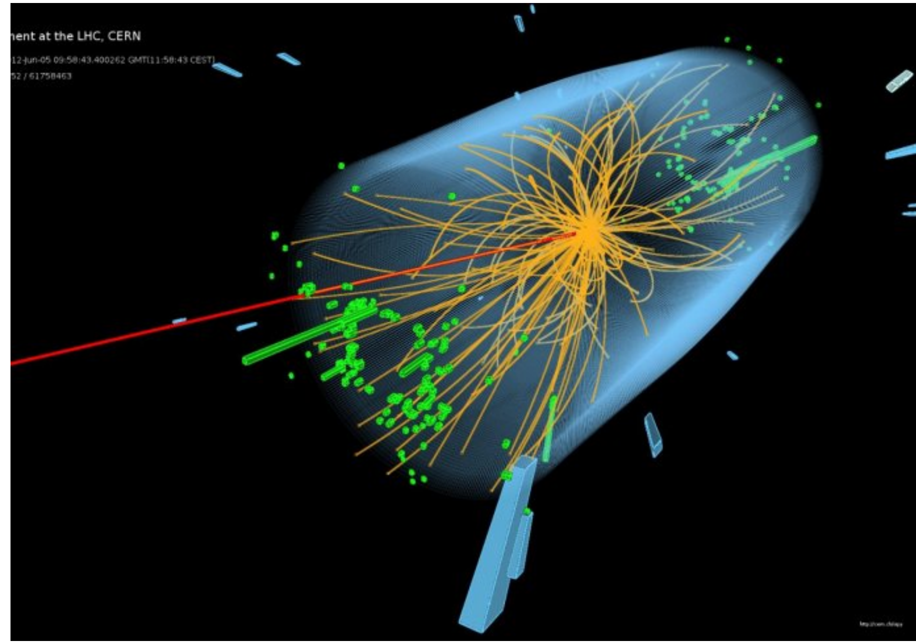
Laboratoire Leprince-Ringuet (LLR)
Ecole Polytechnique - Université Paris-Saclay
CNRS-IN2P3

JRJC 2015

# Introduction

- ATLAS-CMS Higgs couplings measurement combination recently brought evidence of H->$\tau\tau$ decay with a 5.5σ significance (exp. 5.0σ):
  first direct observation of Yukawa couplings!

- Yet no independent discovery by either of the two experiments

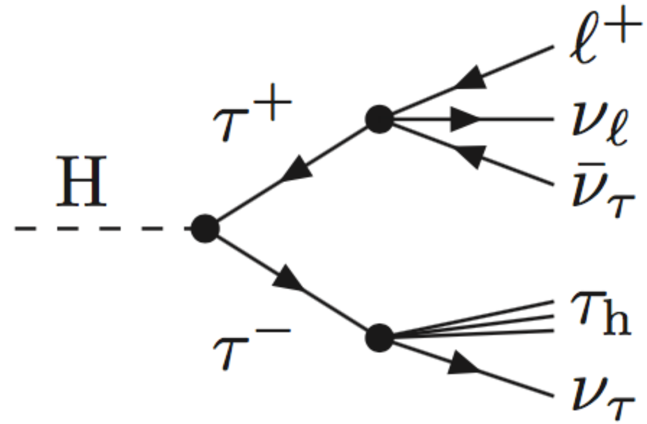


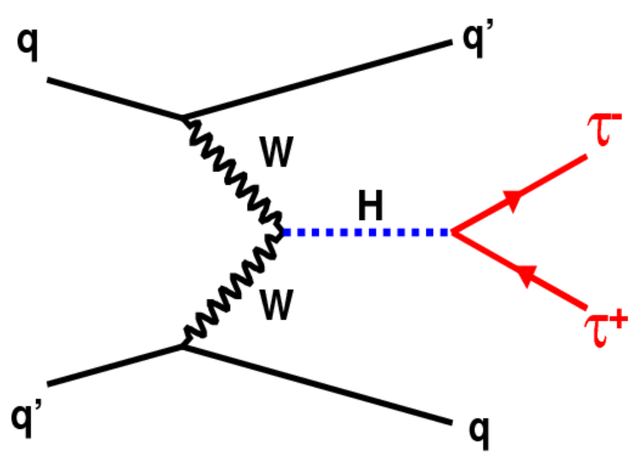
| Channel | References for individual publications | | Signal strength [$\mu$] from results in this paper (Section 5.2) | | Signal significance [$\sigma$] | | |
|---|---|---|---|---|---|---|---|
| | ATLAS | CMS | ATLAS | CMS | ATLAS | CMS | |
| $H \to \tau\tau$ | [58] | [59] | $1.41^{+0.40}_{-0.35}$ | $0.89^{+0.31}_{-0.28}$ | 4.4 | 3.4 | obs. |
| | | | $(^{+0.37}_{-0.33})$ | $(^{+0.31}_{-0.29})$ | (3.3) | (3.7) | exp. |

*CMS-PAS-HIG-15-002*

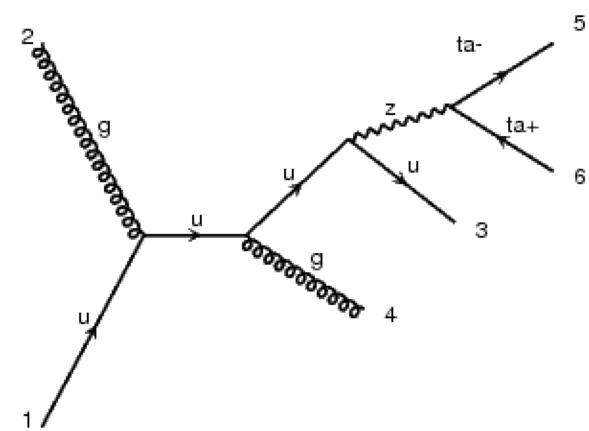
# Introduction

- H->$\tau\tau$ decay challenging to reconstruct because of neutrinos in the final state
  => only charged leptons or hadronic decay products are visible in the detector



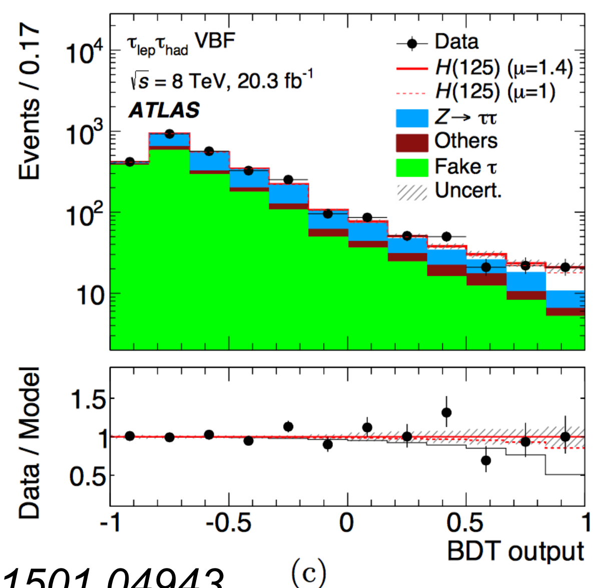




- Sensitivity to H->$\tau\tau$ can be enhanced by looking for associated production like VBF
  => look for two forward jets

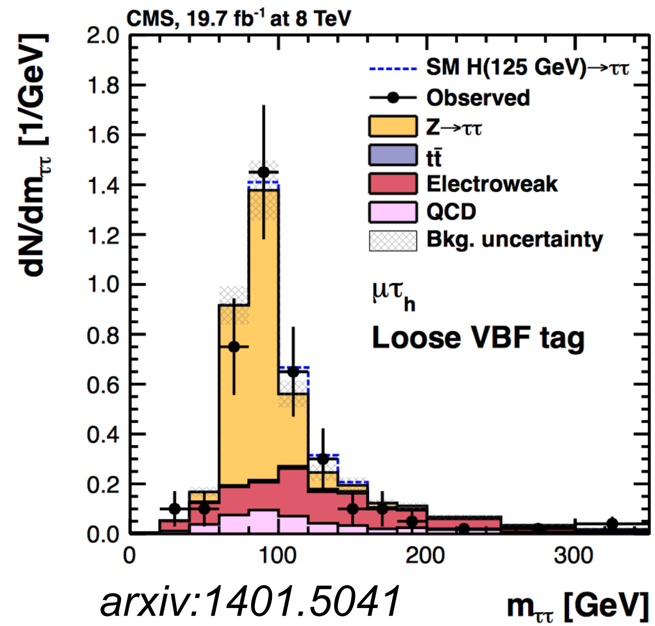


# Introduction

- Di-$\tau$ mass reconstruction based on a likelihood approach in ATLAS (MMC) and CMS (SVFitMass)
- Signal extraction in ATLAS based on Boosted Decision Tree (BDT) while CMS analysis uses directly di-$\tau$ mass estimation




*arxiv:1501.04943*




*arxiv:1401.5041*

- New analysis method for H->$\tau\tau$, based on Matrix Element Method (MEM) will be presented here

# Outline

1. Principles of the Matrix Element Method

2. Application to VBF H->$\tau\tau$

3. Prospects for ttH H->$\tau\tau$

# Principles of the Matrix Element Method

# Principles of the Matrix Element Method

- Goal of the MEM: provide a discriminating variable L(**y**) which optimally combines the available information from the objects reconstructed in the detector (which are associated to a set of observables **y**)

- Final discriminant L(**y**) can be directly fitted or used to define categories with different S/B, thus improving the sensitivity of the search

- Combines theoretical predictions at LO for a given model with information about the detector resolution (possible to take spin-correlation into account)

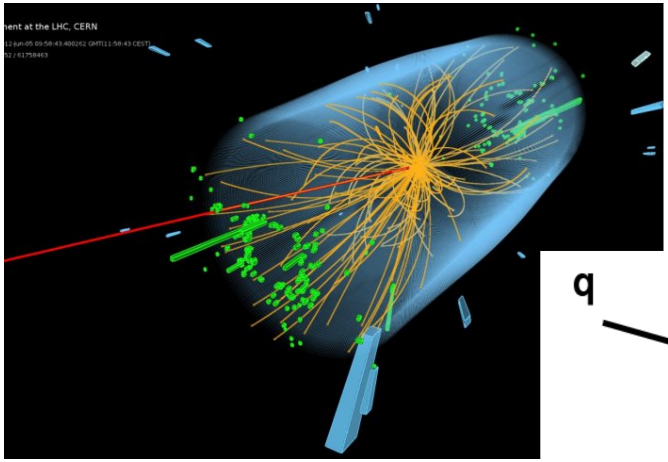
- Requires numerical integration over poorly determined or unmeasured parton quantities (jet energy, neutrinos momenta…)

- Already used in CMS:
  - H->ZZ->4l: MELA, no integration (arxiv:1312.5353)
  - Zbb: cross-check of standard analysis (arxiv:1402.1521)
  - ttH, H->bb: full MEM based analysis (arxiv:1502.02485)

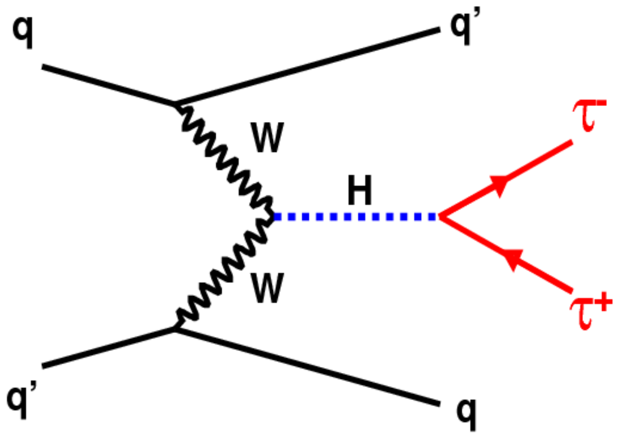
**In this presentation: VBF and ttH, H->$\tau\tau$**

# Principles of the Matrix Element Method

- Observables **y**:
  - leptons
  - $\tau_h$
  - jets
  - MET



- Phase-space point **x**:
  - $\tau$ (before decay)
  - quarks
  - neutrinos




- The final discriminant used is in principle

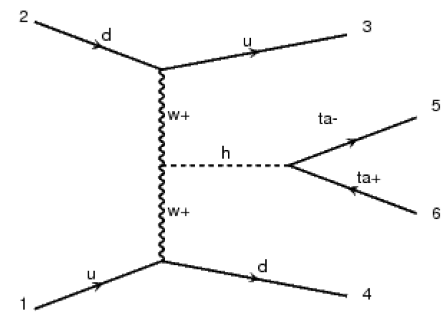
$$\mathcal{L}(\mathbf{y}) = \frac{w_S(\mathbf{y})}{w_S(\mathbf{y}) + w_B(\mathbf{y})}$$

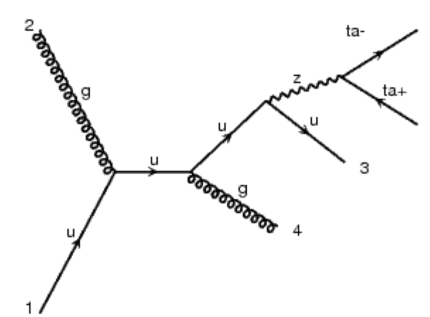
with

$$w_i(\mathbf{y}) = \frac{1}{\sigma_i} \sum_p \int d\mathbf{x} dx_a dx_b \frac{f(x_a, Q) f(x_b, Q)}{x_a x_b s} \delta^2(x_a P_a + x_b P_b - \sum p_k) |\mathcal{M}_i(\mathbf{x})|^2 W(\mathbf{y}||\mathbf{x})$$

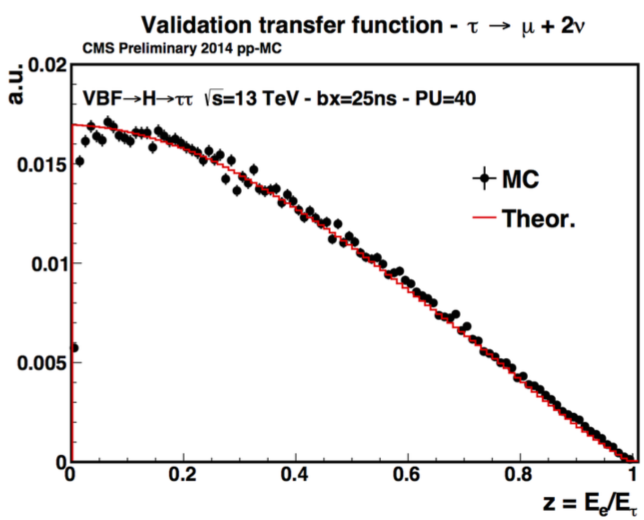
# Principles of the Matrix Element Method

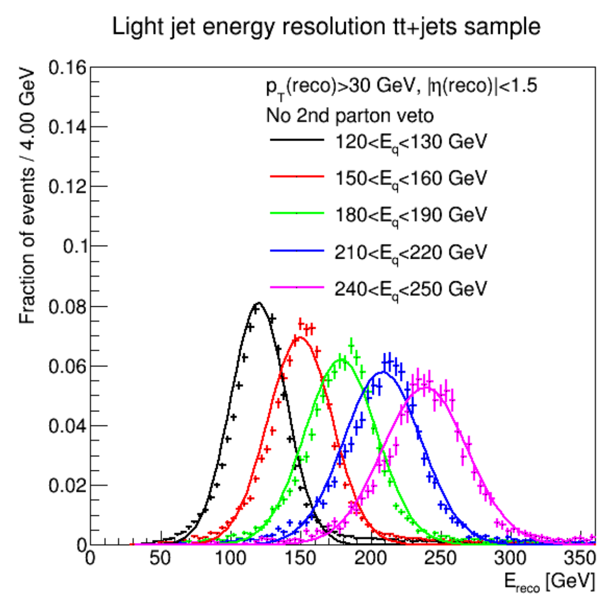
- **Main parts**
  - Matrix Element computed at LO with MadGraph




  - Transfer functions = probability of measuring **y** given a point **x** in the phase space of the final-state particles



# Principles of the Matrix Element Method

- **Monte-Carlo integration**
  - Not possible to compute integral analytically

$$w_i(\mathbf{y}) = \frac{1}{\sigma_i} \sum_p \int d\mathbf{x} dx_a dx_b \frac{f(x_a,Q)f(x_b,Q)}{x_a x_b s} \delta^2(x_a P_a + x_b P_b - \sum p_k) |\mathcal{M}_i(\mathbf{x})|^2 W(\mathbf{y}||\mathbf{x})$$

  - MC integration very efficient to perform numerical integration in multi-dimensional space

The problem Monte Carlo integration addresses is the computation of a multidimensional definite integral

$$I = \int_\Omega f(\overline{\mathbf{x}})\, d\overline{\mathbf{x}}$$

where Ω, a subset of $\mathbf{R}^m$, has volume

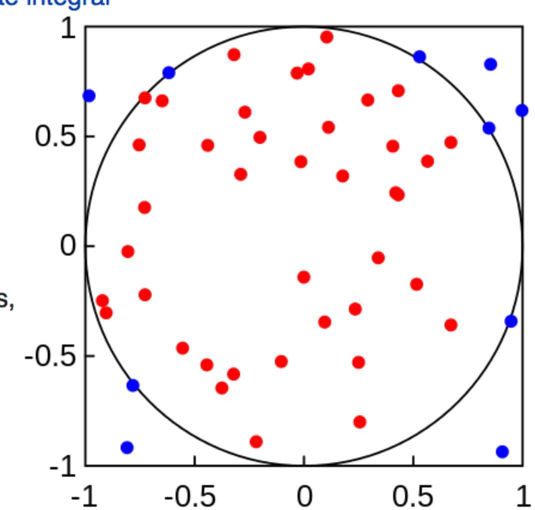
$$V = \int_\Omega d\overline{\mathbf{x}}$$

The naive Monte Carlo approach is to sample points uniformly on Ω:[4] given *N* uniform samples,

$$\overline{\mathbf{x}}_1, \cdots, \overline{\mathbf{x}}_N \in \Omega,$$

*I* can be approximated by

$$I \approx Q_N \equiv V \frac{1}{N} \sum_{i=1}^{N} f(\overline{\mathbf{x}}_i) = V \langle f \rangle.$$




*from https://en.wikipedia.org/wiki/Monte_Carlo_integration*

# Principles of the Matrix Element Method

- **Dimensionality reduction**

$$w_i(\mathbf{y}) = \frac{1}{\sigma_i} \sum_p \int d\mathbf{x} dx_a dx_b \frac{f(x_a, Q) f(x_b, Q)}{x_a x_b s} \delta^2(x_a P_a + x_b P_b - \sum p_k) |\mathcal{M}_i(\mathbf{x})|^2 W(\mathbf{y}||\mathbf{x})$$

- With brute-force computation, number of dimensions is pretty large: for VBF H->$\tau\tau$, up to 8 particles in the final state => 8x3 dimensions
  => very large number of points to perform MC integration

- Function to integrate can be very peaked (signal ME around $m^2_{\tau\bar{\tau}}$ = (125 GeV)$^2$ )

- Possible solution: try to reduce number of dimensions with TF with perfect resolution (direction jets, lepton 3-momentum) + adapted change of variables
  => involved computation to get to the result but possible to go from 24 -> 4 dimensions for VBF H->$\tau\tau$ (=jet energies + $d|\vec{\tau}_l| d\cos\theta_{\tau\bar{\tau}}$ )

# Principles of the Matrix Element Method

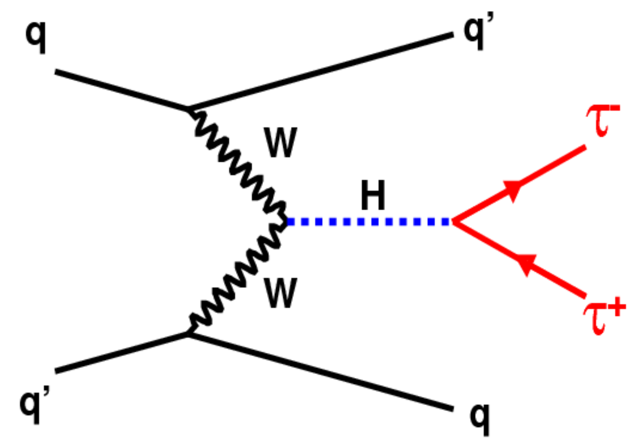
- **Advantages**
  - In principle, optimal combination of theoretical information (matrix element) with detector resolution (transfer function)
  - Can treat complex final states with several relevant observables (jets, di-$\tau$ system, top quarks…), including polarization
  - No training required => not sensitive to input samples (low statistics, inaccurate modelisation…)

- **Drawback**
  - Demanding in terms of computing ressources for MC integration (even after dimensionality reduction)
    => implementation on GPU's currently under development

# Application to VBF $H \rightarrow \tau\tau$

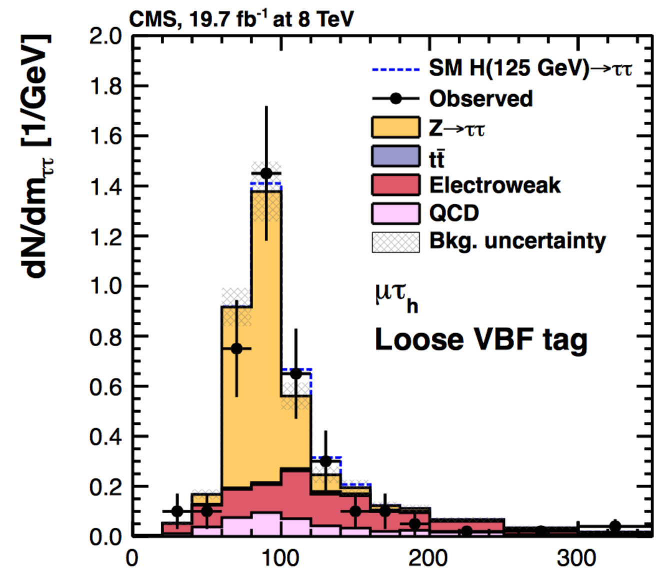
# Application to VBF H->$\tau\tau$

- Validation of the MEM with H->$\tau\tau$ first done for VBF production mode in Run 1 data since CMS analysis already existing and can be used as benchmark




- Standard analysis defines VBF categories based on kinematic of two forward jets and uses SVFitMass distribution for signal extraction

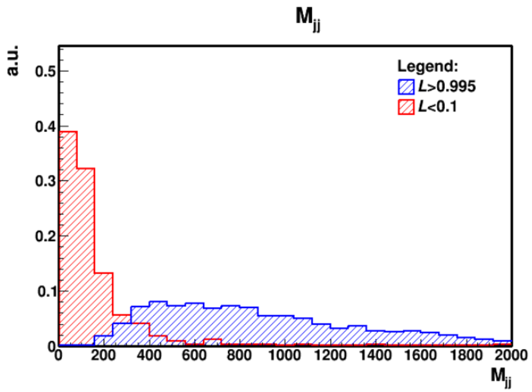
- Main background from Z->$\tau\tau$ (DY+2 jets) production: only one used for now in MEM (ongoing studies to include W+jets background)

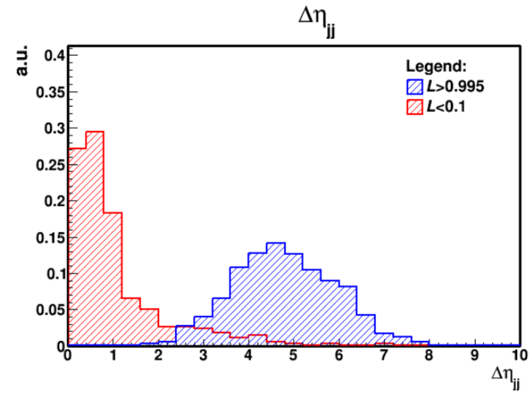



arxiv:1401.5041

# Application to VBF H->$\tau\tau$

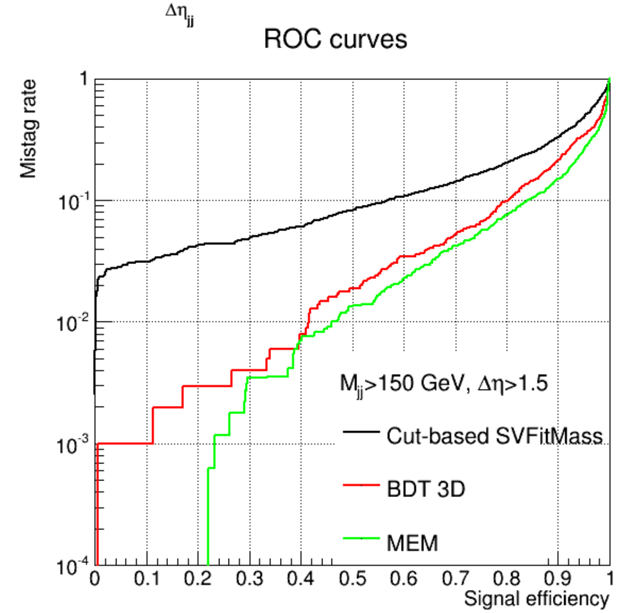
- Since kinematics of the process well-known, possible to do cross-checks to check the correlation between the MEM likelihood ratio and the relevant observables



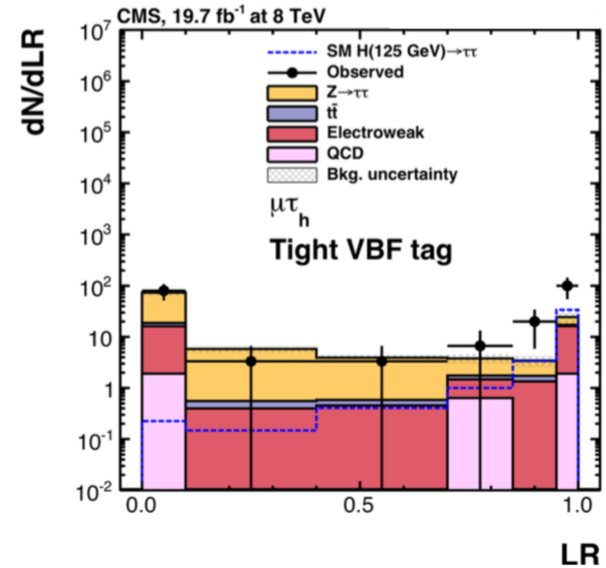

*VBF sample*

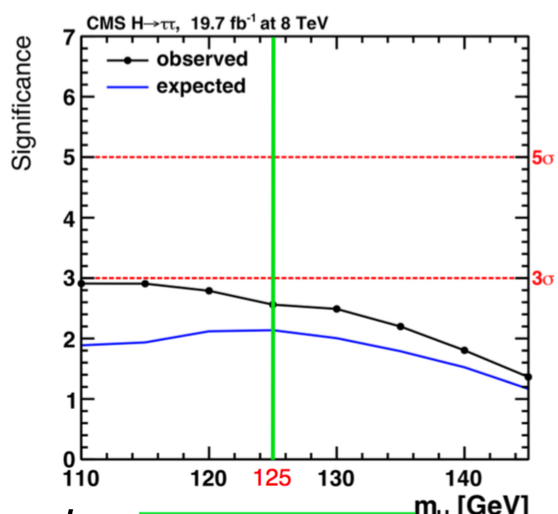
- Performance of the MEM compared to SVFitMass alone and BDT using SVFitMass + Mjj + Δη
  => with current MEM settings, significant improvement with respect to default analysis

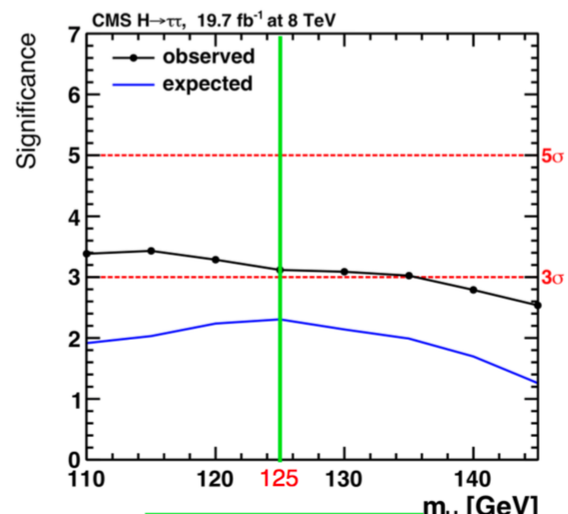


# Application to VBF H->$\tau\tau$

- Full CMS H->$\tau\tau$ analysis set-up reproduced including MEM (only in VBF categories)
- Evaluation of the final limits already performed in $\mu\tau$ channel (*L. Mastrolorenzo's PhD thesis*)
- If improvement confirmed in other channels, can use the MEM for Run 2 analysis and contribute to the discovery of H->$\tau\tau$ independently by CMS




*Analysis w/o MEM*

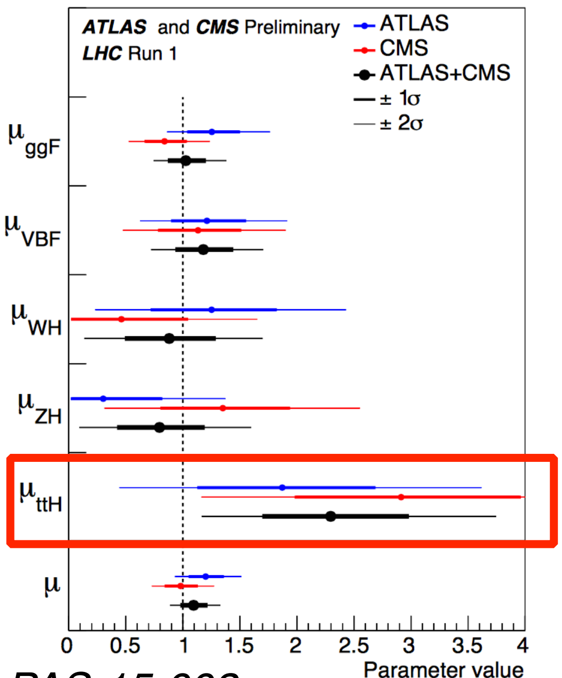
Expected: 2.14
Observed: 2.56

Expected: 2.31
Observed: 3.12

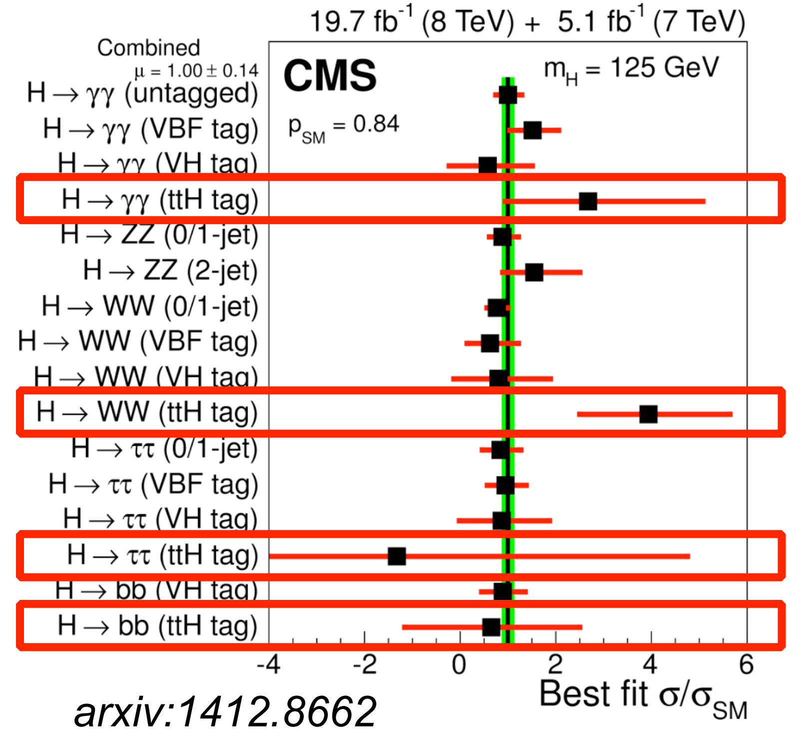
*Analysis with MEM in VBF categories (no MEM in ggF-dominated categories)*

# Prospects for ttH H->$\tau\tau$

# Prospects for ttH H->$\tau\tau$




CMS-PAS-15-002




arxiv:1412.8662

- ATLAS-CMS combination confirmed observed excess in ttH production mode
- Interesting channel to look at for Run 2: direct measurement of top-Higgs coupling
- Among all the ttH channels looked for by CMS at Run 1, H->$\tau\tau$ has the less precise measurement: large margin for improvement with MEM

# Prospects for ttH H->$\tau\tau$

- Presence of $\tau_h$ in final-state can be used to decouple from ttH multi-lepton analysis

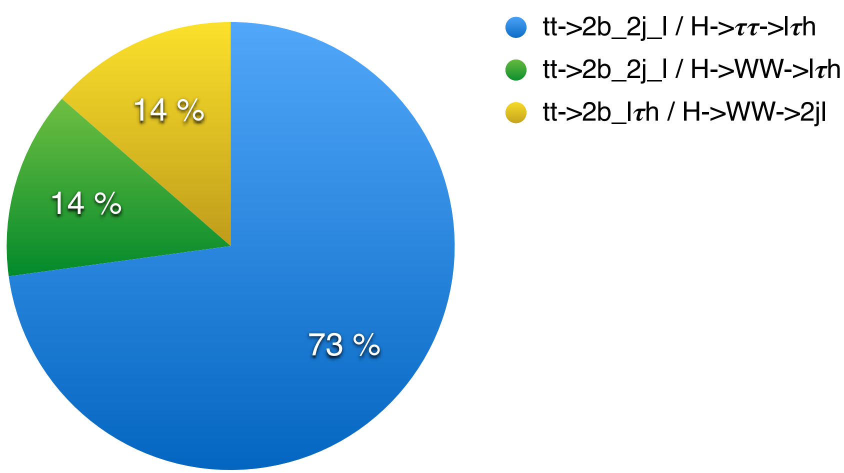
- Multiple leptons in addition can still be used to reduce the background (possibility to require 2l OS / 2l SS / 3l)

- Preselections: ≥ 2 leptons, ≥ 1 $\tau_h$, ≥ 3 jets

- **2 lep SS + 1$\tau_h$ + jets:**

**7.9% ttH, H->$\tau\tau$**

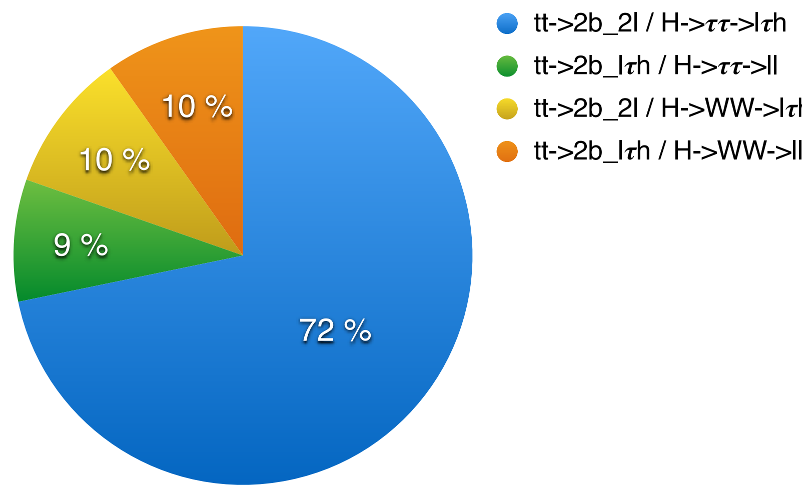
ttH repartition




- **3 lep + 1$\tau_h$ + jets:**

**2.8% ttH, H->$\tau\tau$**

ttH repartition



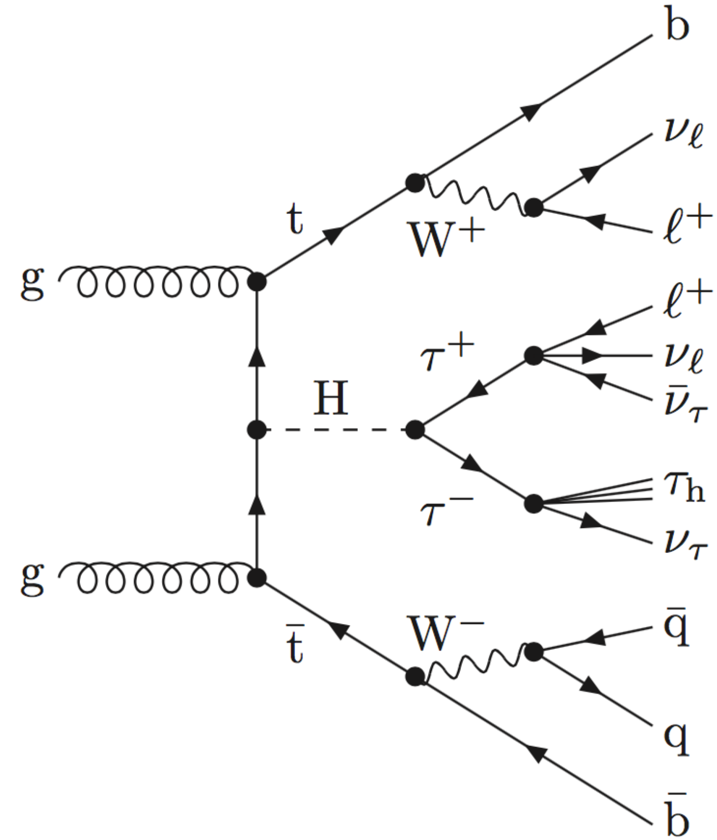
# Prospects for ttH H->$\tau\tau$

- Although final-state more complex than VBF, possibility to rely on b-tagging to identify b-jets + sum over ambiguous permutations with MEM

- Categorization based on b-tagging requirements (2 CSV-medium tagged / 1 CSV-medium + 1 CSV-loose)

- Further categorization based on compatibility of untagged jets with W mass
  => if no W-tagged pair of jets, possibility to integrate over the missing jet direction

- Performance on various backgrounds under study



# Conclusion

- Matrix Element Method is a powerful tool to extract signal, which does not require training and is therefore not impacted by samples with low statistics

- MEM has been successfully implemented for H->$\tau\tau$ analyses, shows very promising results and could lead to the 5σ discovery of H->$\tau\tau$

- Could also provide a more precise measurement of ttH H->$\tau\tau$ and confirm (or not) the excess seen in other ttH channels

# Back-up

# Principles of the Matrix Element Method

$$w_i(\mathbf{y}) = \frac{1}{\sigma_i} \sum_p \int d\mathbf{x} dx_a dx_b \frac{f(x_a, Q) f(x_b, Q)}{x_a x_b s} \delta^2(x_a P_a + x_b P_b - \sum p_k) |\mathcal{M}_i(\mathbf{x})|^2 W(\mathbf{y}||\mathbf{x})$$

| | |
|---|---|
| $\frac{1}{\sigma_i}$ | normalization coefficient (independent of **y**) |
| $\sum_p$ | sum over all the potential associations between the reconstructed objects and the final-state particules + over the potential processes |
| $\int d\mathbf{x} dx_a dx_b$ | integration over the phase space of the final-state particles **x** and over the momentum fractions $x_a$, $x_b$ of the incoming partons |
| $f(x_a, Q) f(x_b, Q)$ | PDFs of the incoming partons, evaluated with LHAPDF |
| $\delta^2(x_a P_a + x_b P_b - \sum p_k)$ | $\delta$-function enforcing the conservation of energy and longitudinal momentum between the incoming partons and the final-state particles |

# Principles of the Matrix Element Method

$$w_i(\mathbf{y}) = \frac{1}{\sigma_i}\sum_p \int d\mathbf{x}dx_a dx_b \frac{f(x_a,Q)f(x_b,Q)}{x_a x_b s}\delta^2(x_a P_a + x_b P_b - \sum p_k)|\mathcal{M}_i(\mathbf{x})|^2 W(\mathbf{y}||\mathbf{x})$$

$|\mathcal{M}_i(\mathbf{x})|^2$ matrix element (ME) squared of the process *i* at LO (for instance ud->udH,H->$\tau\tau$)

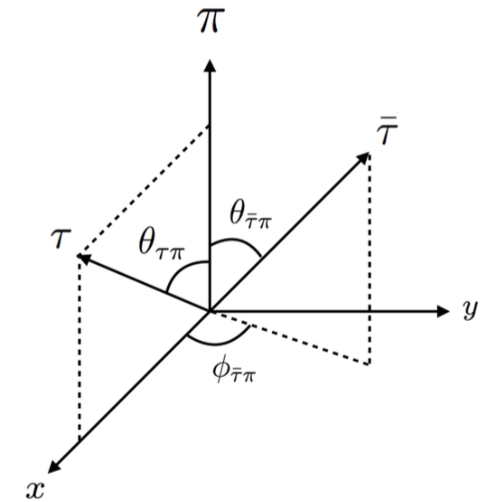
$W(\mathbf{y}||\mathbf{x})$ transfer function = probability of measuring **y** given a point **x** in the phase space of the final-state particles

describes the decay of the unstable final-state particles of the hard-scattering ($\tau$)

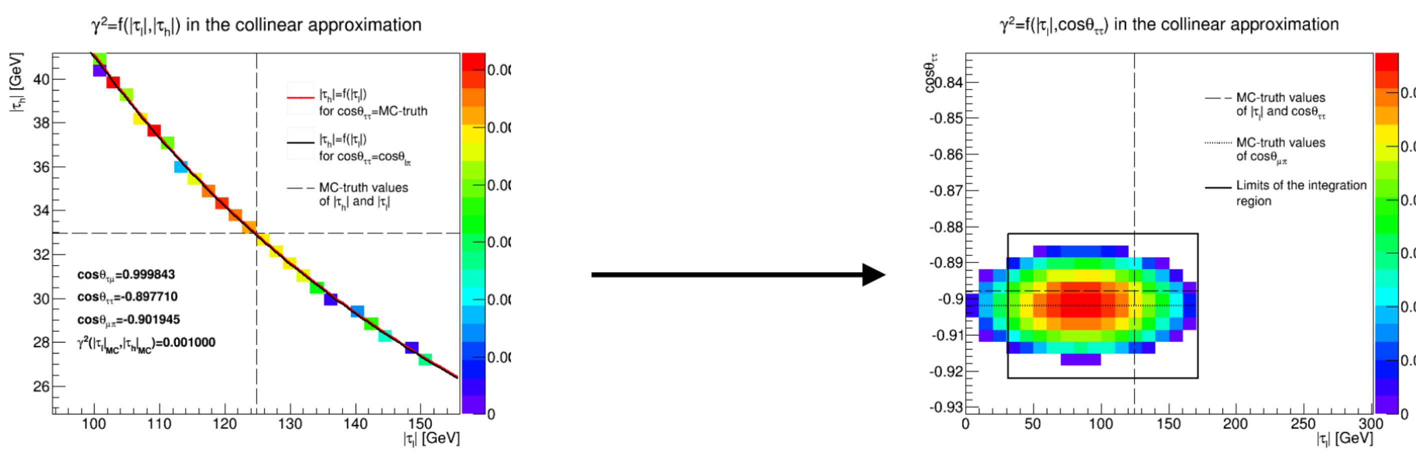
+ takes into account the resolution of the detector on the energy of the jets and on the MET

# Principles of the Matrix Element Method

- **Monte-Carlo integration**
  - Because of mass-shell constraints for $\tau$ + narrow width of the Higgs, change of variables required for the integration
    => complex reconstruction of the di-$\tau$ system from those variables + measured observables



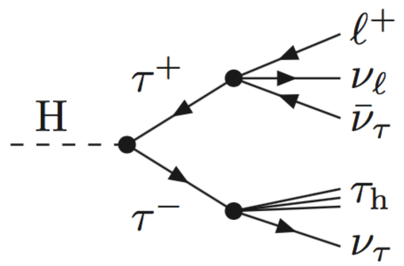

  - Non-trivial kinematic constraints to take into account




$$d|\vec{\tau}_l|d|\vec{\tau}_h|dm^2_{\tau\bar{\tau}}$$

$$d|\vec{\tau}_l|d\cos\theta_{\tau\bar{\tau}}dm^2_{\tau\bar{\tau}}$$

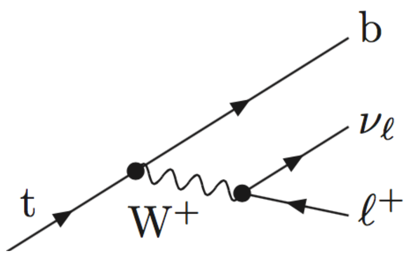
# Principles of the Matrix Element Method




- **Higgs (Z) decay to $\tau\tau$**

  => 2 integrations over $d|\vec{\tau}_l|d\cos\theta_{\tau\bar{\tau}}$

  + optional integration over $dm^2_{\tau\bar{\tau}}$

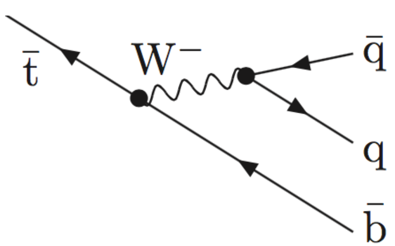



- **Leptonic top decay**

  Enu determined from lepton momentum + $M_W$ constraint

  Eb determined from W momentum + $M_t$ constraint

  => 2 integrations over neutrino direction




- **Hadronic top decay**

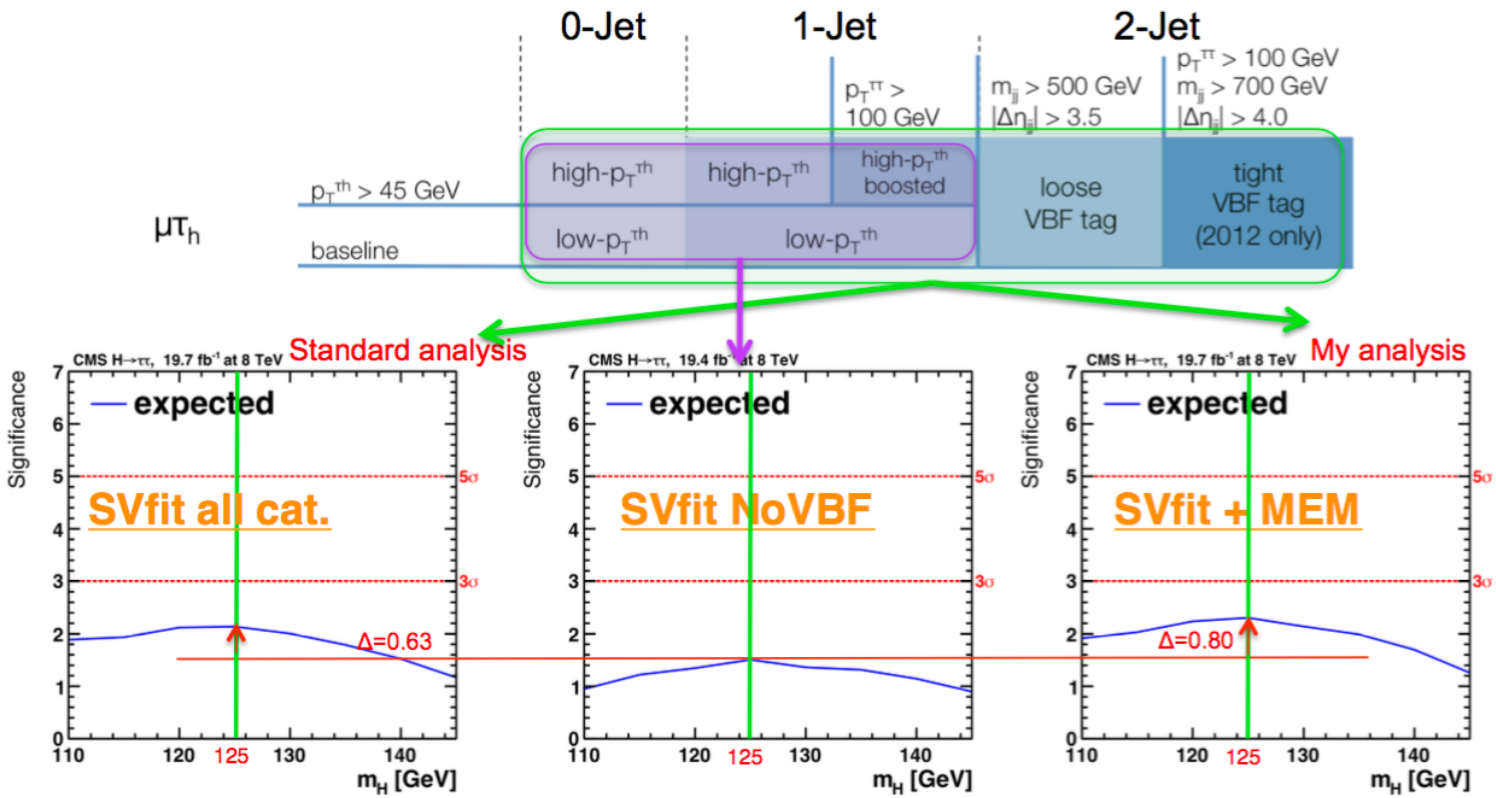
  Eqbar determined from q momentum + $M_W$ constraint

  Eb determined from W momentum + $M_t$ constraint

  => 1 integration over Eq

# Application to VBF H->$\tau\tau$

## Improvement in the VBF only categories

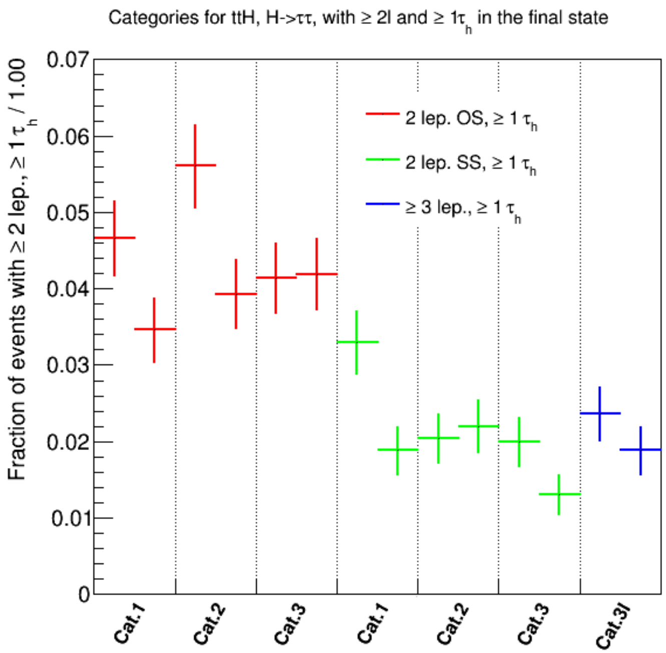



**Improvement in the VBF only categories ~27%!!**

*L. Mastrolorenzo's PhD thesis*

# Prospects for ttH H->$\tau\tau$

- Cat. 1: ≥ 2 leptons, ≥ 1 $\tau_h$, ≥ 4 jets, W-tagged pair of light jets
  Cat. 2: ≥ 2 leptons, ≥ 1 $\tau_h$, ≥ 4 jets, no W-tagged pair of light jets
  Cat. 3: ≥ 2 leptons, ≥ 1 $\tau_h$, 3 jets
- Use of narrow-width approximation for top, W and Higgs keep the number of dimensions for integration as low as possible (VBF H->$\tau\tau$ 4 dim.)
- Performance on various backgrounds under study




*In each category, 2 CSVM / 1 CSVM + 1 CSVL distinction*

| 2 lep. Cat. 1 | 2 lep. Cat. 2 | 2 lep. Cat. 3 | 3 lep. |
|---|---|---|---|
| 5 dim. | 7 dim. | 7 dim. | 6 dim. |